Identidades e tomadas de decisão automatizadas: uma visão crítica sobre os diversos perfis no ecossistema da tecnologia

Marina Regina Arvigo¹

RESUMO

Este artigo tem como pretensão principal colocar em pauta duas questões pouco debatidas até o presente momento sob o ponto de vista acadêmico: (i) a criação de perfis online e sua relação com a autodeterminação informativa, e (ii) a manutenção de um perfil homogêneo de atores participantes nos espaços de criação e tomada de decisão a respeito dos mecanismos tecnológicos. Observa-se, de um lado, que o desenvolvimento, pela indústria, de práticas discriminatórias que perpetuam e promovem preconceitos estruturais não tem sido acompanhado de perto pela produção acadêmica, principalmente no campo jurídico. Por outro lado, a sub-representação de minorias e determinados grupos nos espaços e processos que desenvolvem e consolidam os mecanismos tecnológicos tem sido objeto de variados debates acadêmicos. Porém, a repercussão de demandas de inclusão no ambiente corporativo do ecossistema da tecnologia não tem satisfeito a inquietação de pesquisadores e ativistas no tema. Neste cenário, este artigo propõe que, ainda que os debates sobre estes dois "perfis" no ecossistema da tecnologia tenham caminhado em paralelo até agora, os atores interessados nesses temas tracem uma intersecção possível entre ambos para que, juntos, seja possível propor soluções que viabilizem uma internet menos discriminatória, mais inclusiva e mais justa para todos os usuários da rede.

¹ marina.arvigo@usp.br

INTRODUÇÃO

Assim como ocorre nos passos traçados para o total entendimento sobre o que é e como se estrutura a criação de perfis na internet (a técnica chamada de "*profiling*"), o estabelecimento, desde já, de quais são as camadas de ideias que estruturam o presente artigo é indispensável para a compreensão da tese desenvolvida ao longo das próximas páginas.

No plano de fundo do presente artigo está a discussão sobre a existência, ou não, e a possibilidade, ou não, de haver discriminação algorítmica. Neste ponto, o debate sobre se algoritmos, enquanto criação humana, não existentes por si só e desprovidos de agência, podem ou devem ser responsabilizados por violações a direitos expressos no ordenamento jurídico deve ser apartado para não desviar a rota estabelecida na pesquisa. Assim, deve-se afastar também a incidência ou não da responsabilidade civil sobre o arquiteto das soluções tecnológicas que discriminam, pois, ainda que numa relação de criador e obra explorada mais à frente, tal temática não está abarcada no escopo e finalidade desta redação. O que se pretende afirmar de plano, portanto, é que algoritmos podem, de forma voluntária ou não, perpetuar certos padrões (de comportamento, por exemplo) que, por si só, são discriminatórios. E, por consequência, as estruturas que concebem tais modelos (definidos, aqui, tanto como os resultados do processamento de dados por inteligência artificial, quanto pelos paradigmas pré-definidos e impositivos do ponto de vista social) acabam se legitimando pela retroação do ecossistema em que estão inseridos.

Isto posto, cumpre salientar que a tese central do presente texto é demonstrar que duas das causas - e, portanto, os pontos-chave para a reversão do quadro apresentado acima - da perpetuação dos padrões pelos algoritmos treinados por inteligência artificial e, em última instância, pela manutenção das estruturas discriminatórias, são (i) as tomadas de decisão automatizadas com base em perfis criados por coleta de dados e inferência de padrões, e (ii) a sub representação de minorias políticas nos espaços de criação e tomada de decisão do ecossistema da tecnologia.

Por fim, pretende-se concluir que o enfrentamento da realidade e, por consequência, a construção de uma internet - e, por que não falar, de um ecossistema tecnológico ou, em última instância, um mundo? - mais justa e menos discriminatória passa necessariamente pela construção de uma intersecção entre ambos os campos de batalha.

CONCEITOS E ASPECTOS INTRODUTÓRIOS SOBRE INTELIGÊNCIA ARTIFICIAL

Neste ponto, é necessário estabelecer os patamares mínimos de conhecimento técnico a respeito dos conceitos empregados pela literatura acadêmica nos temas transversais à inteligência artificial².

Assim, esclareço que, para fins deste artigo, deve-se ter em mente, em consonância com o definido por Schermer (2011):

- Inteligência artificial é o conceito usado para descrever sistemas computacionais que são capazes de aprender a partir de suas próprias experiências e resolver problemas complexos em diferentes situações – habilidades que anteriormente pensamos ser únicas em seres humanos. Trata-se, também, de um termo guarda-chuva que engloba diversos tipos de “*machine learning*”, que pode ser definido como “um conjunto de técnicas e mecanismos que permite que computadores ‘pensem’ ao criar algoritmos matemáticos baseados em dados acumulados”.
- Algoritmo, por sua vez, é uma sequência finita de instruções bem definidas e não ambíguas, cujas características são: (i) finitude (de passos, tempo e recursos), (ii) exatidão (clareza e precisão), e (iii) efetividade. Neste sentido, quanto maior a complexidade da tarefa a ser realizada, maior deve ser o

² Considerando que o domínio de termos técnicos do ecossistema tecnológico ainda não é uniforme em todos os campos dos saberes, demonstra-se mais acertada a escolha pela determinação das definições empregadas ao longo do presente artigo. Além disso, parto de um pressuposto que, a meu ver, deveria ser compartilhado entre pesquisadores, independentemente do seu objeto de estudo ou da área do conhecimento sobre o qual se debruçam: um debate honesto, principalmente sobre inclusão, como ocorre no presente caso, deve dar condições que permitam a indivíduos que não pertencem a um grupo privilegiado do ponto de vista informacional (ou, resumindo em uma palavra, “*outsiders*”) não somente entenderem o que está escrito, mas também se inserirem e debaterem o saber compartilhado pelas palavras que, juntas, expressam sua pesquisa.

detalhamento das instruções a fim de evitar ambiguidade. Também vale citar a conceituação a seguir (LETOUZÉ, 2015 *apud* Cetic.br, 2018): "Na ciência da computação, um algoritmo é uma série de instruções ou regras predefinidas, escritas em uma linguagem de programação destinada para dizer a um computador como resolver sequencialmente um problema recorrente, especialmente ao envolver a realização de cálculos e o processamento de dados. Há um crescente uso de algoritmos para fins de tomada de decisão em uma progressiva gama de atividades e indústrias, tais como policiamento e serviços bancários".

- "*Machine learning*", por outro lado, engloba o conceito de "*deep learning*", considerando-se que alguns tipos de "*deep learning*" são criados com os mesmos princípios de redes neurais do cérebro e geralmente são baseados em um conjunto conhecido de dados de treinamento ("*training data*") que auxilia os algoritmos de auto aprendizado a desenvolver uma tarefa. Também fazendo referência a LETOUZÉ (*apud* Cetic.br, 2018): "[aprendizado de máquina] refere-se à construção e estudo de algoritmos computacionais – procedimentos, passo a passo, para cálculos e/ou classificação – que podem aprender a crescer e mudar quando expostos a novos dados. Representa a capacidade das máquinas de 'aprender' a fazer previsões e tomar decisões melhores com base em experiências passadas, como na filtragem de *spam*, por exemplo".

Há dois aspectos principais da inteligência artificial de especial relevância para a privacidade: (i) a possibilidade do software poder tomar decisões, e (ii) o sistema se desenvolver ao aprender com sua própria experiência.

Um dos impasses que surgem a partir de ambos os aspectos é a "*black box*", termo que, ao fazer referência à dificuldade que se apresenta em determinados momentos em saber como um resultado é produzido na "*machine learning*", também se conecta ao conceito de opacidade no tratamento de dados pessoais e, conseqüentemente, à obscuridade no processo de tomadas de decisão automatizadas.

Independentemente de que tipo de algoritmo ou método utilizado pela “*machine learning*”, o resultado será um modelo, um termo guarda-chuva para “*machine learning*”. Quando falamos de modelos, a facilidade com a qual os resultados podem ser checados varia bastante, de modo que se pretende demonstrar, em uma escala de fácil/difícil, como é possível checar o caminho deles.

Um dos modelos mais simples é a “*decision tree*”, cujo funcionamento se dá da seguinte forma: uma pessoa começa pelo topo e cada nível seleciona uma resposta baseada em um valor de uma determinada característica. Essa pessoa continua até a base da árvore, onde o resultado final (a decisão) é encontrado. Este tipo de modelo fornece um alto grau de transparência, pelo menos quando a árvore é baseada em uma quantidade viável de dados. É possível mover para cima pela árvore para ver o critério na qual o resultado é baseado.

“*Neural networks*” são usados em uma metodologia que é em muito inspirada pelo nosso entendimento da forma como o cérebro funciona. Tais redes são construídas pelo que é um componente muito simples – o “*perceptron*” – que, se engatilhados uns aos outros, podem criar redes grandes e complexas.

Cada variável tem um peso que determina o quanto a característica presente no input respectivo influenciará no resultado final.

Uma rede neural consiste em três partes: “*input layer*”, “*hidden layers*”, e “*output layer*”. Se há mais de um “*hidden layer*”, então o “*machine learning*” é considerado um “*deep learning*”.

Um dos desafios das redes neurais é que, da forma como mostrados acima, cada “*input*” é visto de forma isolada. Ocorre que, em muitas situações, trabalha-se com informações que têm um contexto. Por exemplo, algumas palavras possuem um significado diferente a depender do contexto em que estão inseridas, e este contexto nem sempre é expresso pela mesma sentença. É também por isso que algumas redes neurais possuem a forma de uma memória de curto tempo (“*short-term memory*”). Isso tanto permite diferentes “*outputs*” baseados nos “*inputs*” inseridos previamente, quanto significa que pode ser muito difícil simplesmente

examinar os algoritmos para entender como eles funcionam e quais decisões eles tomam.

OS "PERFIS" NO ECOSSISTEMA DA TECNOLOGIA

A Técnica de Profiling

O Que é “profiling”?

De acordo com a literatura acadêmica, a técnica de *profiling* pode ser definida como:

“Profiling is the process of discovering correlations between data in databases that can be used to identify and represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category”³.

Desde táticas antiterrorismo⁴, passando por microdirecionamento⁵ e impulsionamento de conteúdo em época eleitoral⁶, *credit scoring* até chegar à

³ Hildebrandt, M. (2008): Defining Profiling: A New Type of Knowledge? In: *Profiling the European Citizen, Cross-Disciplinary Perspectives* (Hildebrandt, M., Gutwirth, S., eds.), Springer Science, p. 17.

⁴ “In the wake of the 9/11 terror attacks against the United States, John Poindexter, the highest-ranking official to be criminally convicted in the Iran-Contra debacle in the mid-1980s, successfully pitched an idea to the U.S. Defense Department to create a computerized surveillance system to track the everyday activities of all American citizens with the goal of ferreting out terrorists. The program, known originally as Total Information Awareness (TIA) would apply sophisticated computer data-mining techniques to sift through virtual mountains of data of everyday transactions, such as credit card purchases, e-mail and travel itineraries, in an attempt to discover patterns predictive of terrorist activity. If implemented, TIA's exercise in data mining would, in the words of one critic, ‘amount to a picture of your life so complete it's equivalent to somebody following you around all day with a video camera’.” Andrew J. McClurg, A Thousand Words Are Worth a Picture: A Privacy Tort Response to Consumer Data Profiling, 98 Nw. U. L. Rev. 63 (2003). Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1628724. Acesso em: 20.02.2019.

⁵ “Online political microtargeting involves ‘creating finely honed messages targeted at narrow categories of voters’ based on data analysis ‘garnered from individuals’ demographic characteristics and consumer and lifestyle habits’. Online political microtargeting can take the ‘form of political direct marketing in which political actors target personalized messages to individual voters by applying predictive modelling techniques to massive troves of voter data’. Online political microtargeting could also be seen as a type of behavioural advertising, namely political behavioural advertising. Behavioural advertising is a modern marketing technique that involves tracking people’s online behaviour to use the collected information to display individually targeted advertisements.

publicidade, a técnica de perfilamento⁹ está mais presente no cotidiano dos cidadãos do que se pode imaginar. A cada site que se abre e se lê a frase “este site utiliza cookies⁷ para personalizar sua navegação” ou semelhantes, estamos diante de tal técnica.

Os Riscos Associados à Técnica de profiling: uma crítica ao “viés” de privacidade no olhar sobre o profiling

Conforme aponta Schermer (2011), há atualmente uma supervalorização dos riscos associados entre a técnica de "profiling" e a privacidade informacional dos usuários da rede. O excesso de enfoque dado a esta perspectiva de análise acaba ofuscando os riscos reais que a criação de perfis e a mineração de dados⁸ podem representar para grupos e indivíduos. Os riscos mais significativos associados à criação de perfis e à mineração de dados são:

1. Discriminação: A classificação e divisão estão no centro da mineração de dados (preditiva). Como tal, a discriminação é parte integrante do perfil e da mineração de dados. No entanto, há situações em que a discriminação é considerada antiética e até ilegal. Isso pode ocorrer, por exemplo, quando um

Online political microtargeting is used, for example, to identify voters who are likely to vote for a specific party and therefore can be targeted with mobilising messages. (For ease of reading, we also refer to ‘microtargeting’). Microtargeting also enables a political party to select policy stances that match the interests of the targeted voter – for instance family aid for families, or student benefits for students." Borgesiues et al: Online Political Microtargeting: Promises and Threats for Democracy, Utrecht Law Review, v. 14, issue 1, 2018. Disponível em: <https://www.ivir.nl/publicaties/download/UtrechtLawReview.pdf>. Acesso em: 20.02.2019. Para mais informações a respeito, sugiro a leitura do conteúdo disponível em <http://www.internetlab.org.br/pt/tag/microdirecionamento/>. Também é interessante visitar o site a seguir, cujo projeto monitora o microdirecionamento e perfilação psicológica durante as eleições em diversos países do mundo: <https://whotargets.me>.

⁶ Para mais informações a respeito, sugiro ler “O que esperar do patrocínio de posts de candidatas durante a eleição” (VENTURINI, 2018). Disponível em: <https://www.nexojornal.com.br/expresso/2018/01/07/O-que-esperar-do-patrocinio-de-posts-de-candidatas-durante-a-eleicao>. Acesso em: 20.02.2019.

⁷ “Cookies são arquivos criados pelos websites que você visita. Eles tornam sua experiência on-line mais fácil, economizando informações de navegação. Com os cookies, os sites podem manter você conectado, lembrar suas preferências do site e fornecer conteúdo relevante localmente. Existem dois tipos de cookies: (i) Cookies primários: criados pelo site que você acessa. O site é exibido na barra de endereços; e (ii) Cookies de terceiros: criados por outros sites. Esses sites possuem uma parte do conteúdo, como anúncios ou imagens, que você vê na página da Web que acessa.” Disponível em: <https://support.google.com/chrome/answer/95647?co=GENIE.Platform%3DDesktop&hl=pt-BR>. Acesso em 20.02.2019.

⁸ “Data mining is the process of discovering interesting patterns and knowledge from large amounts of data. The data sources can include databases, data warehouses, the Web, other information repositories, or data that are streamed into the system dynamically” (HAN, KAMBER & PEI, 2012).

exercício de mineração de dados é focado em características como etnia, gênero, religião ou preferência sexual. Mas mesmo sem um desejo prévio de julgar as pessoas com base em características particulares, existe o risco de, inadvertidamente, discriminar determinados grupos ou indivíduos. A razão para isso é que os algoritmos preditivos de mineração de dados podem “aprender” a discriminar com base em dados tendenciosos usados para treinar o algoritmo.

2. Desindividualização: Em muitos casos, a mineração de dados é, em grande parte, relacionada à classificação e, portanto, há o risco de que as pessoas sejam julgadas com base nas características do grupo e não em suas próprias características e méritos individuais. Os perfis de grupo geralmente contêm estatísticas e, portanto, as características dos perfis de grupo podem ser válidas para o grupo e para indivíduos como membros desse grupo, embora não para indivíduos como tal.
3. Assimetrias de informação: A mineração de dados pode levar a informações valiosas para as partes que a empregam. Quando a mineração de dados visa obter mais informações sobre indivíduos ou grupos, encontramos o problema da assimetria de informações. As assimetrias de informação podem influenciar a igualdade de condições entre o governo e os cidadãos, e entre empresas e consumidores, perturbando o atual equilíbrio de poder entre diferentes partes.

A questão sobre a qual este artigo se debruça é o primeiro problema listado por Schermer, sem, no entanto, desprezar os riscos associados às demais problemáticas levantadas pelo autor.

As Três Camadas de Dados dos Perfis na Internet

Szymielewicz (2019) defende a tese, com a qual este artigo está alinhado, de que a identidade digital dos usuários da rede possui três camadas, das quais é possível

proteger, a priori, somente uma. A autora discorre que as camadas que compõem os perfis são⁹:

1. O que você compartilha: Trata-se dos dados que você escolhe compartilhar nas redes sociais e aplicativos móveis. Em tradução autorizada por Szymielewicz¹⁰, diz-se que "[são] dados que você alimenta em mídias sociais e aplicativos móveis. Isso inclui o que você revelou em suas informações de perfil, suas postagens públicas e mensagens privadas, curtidas, consultas de pesquisa, fotos enviadas, testes e enquetes que você realizou, eventos que participou, sites visitados e outros tipos de interações conscientes".
2. O que seu comportamento diz aos algoritmos: São os dados cuja contextualização permite aos algoritmos correlacionarem metadados e, a partir deles, criar padrões do seu comportamento na rede e fora dela.
3. O que a máquina pensa sobre você: É a tomada de decisão automatizada efetivamente realizada ou, em outras palavras, a inferência de conclusões sobre seus dados - sejam eles compartilhados (in)voluntariamente por você (primeira camada), sejam eles o rastreamento de padrões sobre o qual você não tem escolha (segunda camada).

Conforme será demonstrado a seguir, a problemática envolvendo a ausência de controle dos sujeitos de dados sobre as inferências de padrões comportamentais a partir de metadados (segunda camada) e conclusões e interpretações a partir das camadas anteriores (terceira camada) é um debate a ser travado do ponto de vista social e jurídico para que a rede possa ser um ambiente menos anti discriminatório e, por consequência, mais justo para todos os usuários que dela desfrutam.

⁹ Um gráfico, criado pela Panoptikon Foundation (Polônia), que demonstra de forma lúdica quais dados estão associados a cada uma dessas camadas está disponível em: <https://panoptikon.org/sites/default/files/3levels.png>. Acesso em: 20.02.2019.

¹⁰ "Sua identidade digital tem três camadas e você só pode proteger uma delas". CARIBÉ, João Carlos Rebello; MONTEIRO, Renato Leite. Publicado em 01.03.2019, disponível em: <http://pesquisa.wazushi.com.br/2019/03/01/sua-identidade-digital-tem-tres-camadas-e-voce-so-pode-protetger-uma-delas/>. Acesso em: 02.03.2019.

A "Black-Box" e o Direito à Autodeterminação Informativa

O ponto de atenção do presente tópico é a correlação entre as inferências realizadas por algoritmos nas segunda e terceira camadas e a retomada do controle, por parte do sujeito de dados, do que é feito com seus dados e metadados, sendo certo que o conceito de autodeterminação informativa deve extrapolar a noção do controle sobre o que se publica ou não (primeira camada) para também abarcar os padrões de comportamento (segunda camada) e conseqüentes tomadas de decisão automatizadas (terceira camada).

Nas palavras de Doneda (2000), o conceito de direito à autodeterminação informativa reflete:

"Considerando-se a esfera privada como um conjunto de ações, comportamento, preferências, opiniões e comportamentos sobre os quais o interessado pretende manter um controle exclusivo, esta tutela [da privacidade] há de basear-se em um novo "direito à autodeterminação informativa", hoje possível de ser identificado em diversos ordenamentos, que estabelece condições para um efetivo controle das informações pessoais em circulação¹¹."

Assim, nota-se que o direito à autodeterminação informativa, cuja declaração como direito fundamental remonta a 1983 com o julgamento pelo Tribunal Constitucional Federal da Alemanha da Lei do Censo ("*Volkszählungsgesetz*")¹², extrapola a noção de um direito que somente poderia ser exercido no momento de fornecimento dos dados pessoais para uma concepção que leva em consideração um processo

¹¹ DONEDA, Danilo. "Considerações iniciais sobre os bancos de dados informatizados e o direito à privacidade". Disponível em: <http://www.egov.ufsc.br/portal/sites/default/files/anexos/8196-8195-1-PB.htm>. Acesso em: 20.02.2019.

¹² Uma referência interessante para se aprofundar no precedente alemão é a coletânea "Decisões anotadas sobre direitos fundamentais", selecionada e elaboradas pelo próprio Tribunal Constitucional Federal Alemão. A Corte Alemã dedica parte de seu volume I à decisão em questão. Disponível em: https://www.kas.de/c/document_library/get_file?uuid=4f4eb811-9fa5-baeb-c4ce-996458b70230&groupId=252038. Acesso em: 20.02.2019.

contínuo de exercício deste direito mediante o controle sobre a legitimidade do recolhimento (primeira camada), da divulgação (primeira camada) e da utilização - ou, em palavras mais contemporâneas, o tratamento - dos dados pessoais (segunda e terceira camadas).

No contexto contemporâneo de o que alguns autores chamam de capitalismo de vigilância¹³, é importante pontuar que o conceito de opacidade no tratamento de dados pessoais e consequente obscuridade no processo decisório de tomada de decisões automatizadas é uma oposição de nomenclatura ao direito de transparência e à explicação cristalizados no ordenamento jurídico brasileiro por meio da Lei Geral de Proteção de Dados Pessoais (IGARAPÉ, 2018).

Isto porque, retomando o conceito de "*black box*" explicitado no início deste artigo, o emprego massivo de inteligência artificial e a utilização de "*training data*" em "*machine learning*" torna praticamente impossível explicar como a informação é correlacionada e sopesada em um determinado processo¹⁴.

Neste sentido, para que o sujeito de dados tenha o efetivo controle sobre seus dados, é necessária a proposição de um debate que alcance não somente a escolha do sujeito de dados por disponibilizar, principalmente na rede, suas informações e inclua sob seu guarda-chuva também as observações comportamentais através de metadados e o processo de tomada de decisão automatizada com base nas primeira e segunda camadas dos .

- *TOMADAS DE DECISÃO BASEADAS NAS SEGUNDA E TERCEIRA CAMADAS E IMPLICAÇÕES SOBRE DISCRIMINAÇÃO A PARTIR DO "PROFILING"*

¹³ Cite-se, notadamente, Shoshana Zuboff e seu livro "The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, Public Affairs, 2019. É possível antever parte do debate travado no livro por meio do seguinte artigo: <https://diplomatie.org.br/um-capitalismo-de-vigilancia/>. Acesso em: 20.02.2019.

¹⁴ Segundo relata a Autoridade Norueguesa de Proteção de Dados Pessoais em seu relatório publicado em Janeiro de 2018, outro desafio relacionado às caixas-pretas de algoritmos diz respeito à relação entre revelar informações sobre o modelo para informar sobre o processamento dos dados do sujeito de dados e, possível e consequentemente, revelar também segredos comerciais e direitos de propriedade intelectual. Disponível em: <https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf>. Acesso em: 20.02.2019.

A necessidade de ampliação o escopo da autodeterminação informativa para que se imponha a transparência necessária ao tratamento de dados e à obscuridade no processo de tomada de decisão também encontra sustentação sob o ponto de vista do direito à antidiscriminação.

Tais opacidade e obscuridade mantêm longe do escrutínio público - ou, caso se discuta sobre a relevância da observação pública, que se afirme pelo menos a análise do sujeito de dados - o entendimento a respeito de quais informações têm sido sopesadas com auferimento de maior importância e quais têm tido menor valor nesta ponderação, sendo certo somente que este ambiente é propício para a realização de práticas discriminatórias que perpetuam e promovem problemas e preconceitos estruturais da sociedade contemporânea.

Em síntese, a ausência de transparência e controle sobre as segunda e terceira camadas da identidade digital dos usuários da rede favorece a perpetuação e promoção da discriminação (que estruturam) na sociedade atual.

A Ausência de Diversidade no Perfil dos Atores nos Espaços de Criação e Tomada de Decisão dos Mecanismos Tecnológicos

A Relação Criador e Obra no Ecosistema da Tecnologia e as Implicações Sobre a Tomada de Decisão Automatizadas e Consequente Perpetuação de estruturas e Padrões Discriminatórios

O ponto de partida cujo estabelecimento é necessário para a continuidade do raciocínio do presente artigo trata da sub-representação de minorias no ecossistema da tecnologia, sobretudo nos espaços próprios para a tomada de decisão a respeito de diretrizes de desenvolvimento da tecnologia na sociedade contemporânea. Conforme discorre Saboya (2013):

"Segundo pesquisas realizadas por essa mesma autora (ROSSITER, 1982, 1995, 2003), os preconceitos se revelam na alocação de postos de trabalho: delegam-se às mulheres

tarefas repetitivas e consideradas femininas, como por exemplo, as que demandariam qualificações específicas – expressão essa que seria um eufemismo para tarefas repetitivas e conseqüentemente mal pagas – que exigiriam maior cuidado e atenção (como as relacionadas ao posto de auxiliar nos laboratórios) e que, por conseguinte, deixariam as mulheres fora dos círculos de decisão; em outras palavras, atividades que as impossibilitariam de subir na carreira acompanhando seus colegas homens. Além disso, as mulheres fariam carreiras mais longas, demorando-se mais que os homens nos diferentes níveis, em razão de casamento e/ou filhos, o que lhes exigiria uma dupla jornada de trabalho"¹⁵.

Dito isso, afirme-se que o que se estabelece por meio do presente tópico é que os mecanismos tecnológicos, ainda que não em toda sua extensão¹⁶, mas pelo menos em sua grande medida, acabam incorporando às suas linhas de código e ao processo de tomada de decisão um olhar e entendimento de mundo próprio de seu autor - ou quem estiver no topo de hierarquia que possa sobre o viés , diferenciação que não possui efeitos práticos para a relação que estabeleço entre criador e obra.

Desde a seleção de quais dados são relevantes para que se direcione o alvo do mecanismo, passando pelo dimensionamento do valor dado aos diferentes dados

¹⁵ SABOYA, Maria Clara Lopes (2013): "Relações de Gênero, Ciência e Tecnologia: Uma Revisão da Bibliografia Nacional e Internacional". Educação, Gestão e Sociedade: revista da Faculdade Eça de Queirós, ISSN 2179-9636, Ano 3, número 12, novembro de 2013. Disponível em: <http://uniesp.edu.br/sites/biblioteca/revistas/20170509155548.pdf>. Acesso em: 20.02.2019.

¹⁶ Conforme afirma PASQUALE (2015, p. 6): "Desconstruir as caixas pretas de Big Data não é fácil. Mesmo se estivessem dispostos a expor seus métodos ao público, os modernos setores bancário e da Internet representam difíceis desafios para nosso entendimento desses métodos. As conclusões a que chegam - sobre a produtividade dos funcionários, a relevância dos sites ou a atratividade dos investimentos - são determinadas por fórmulas complexas concebidas por legiões de engenheiros e guardadas por uma falange de advogados" (tradução livre). Cite-se também a publicação do El País: "A perigosa caixa preta dos algoritmos e a campanha eleitoral de 2018" (2017): "Mais recentemente, algoritmos de aprendizagem automática passaram a escrever, sozinhos, outros algoritmos por meio de inteligência artificial, o que, por vezes, pode levar a resultados totalmente inesperados, que não poderiam ser antevistos pelos humanos que desenvolveram o código original. Esse código é, pela maioria das legislações do mundo, proprietário. Isso significa que ele pertence a uma empresa, pode ter um grande valor de mercado e ser considerado um segredo de negócio. O acesso a ele por terceiros pode significar uma grande desvantagem competitiva. E aqui reside um dos maiores embates que impede a efetiva transparência dos algoritmos."

coletados e pelas inferências de padrão possíveis a partir de tais dados, até chegar à tomada de decisão em si, o mecanismo tecnológico não passa de um arquétipo idealizado por seu criador.

Nas palavras da publicação "Panorama Setorial da Internet" (Ano 10, N. 2) do Cetic.br:

"Algoritmos são um reflexo do comportamento humano; até quando não expressos explicitamente, certos vieses presentes nos dados e nas decisões de projeto podem ser influentes no comportamento do sistema. Por isso, em alguns casos, decisões baseadas em algoritmos podem reproduzir ou reforçar um viés negativo, com padrões de discriminação e manutenção de estereótipos, como herança de decisões preconceituosas de pessoas ou simplesmente pelos dados refletirem aspectos culturais, históricos e socio-demográficos existentes na sociedade"¹⁷.

Um exemplo didático para ilustrar a relação entre criador e obra no campo da tecnologia é o site "*This Person Does Not Exist*"¹⁸, por trás do qual existe uma inteligência artificial treinada com milhões de fotos que gera rostos aleatórios que são geralmente super normais. No entanto, o que em um primeiro momento pode "parecer muito Black Mirror", agrega em um só caso os 4 apontamentos mais importantes levantados pela Casa Branca¹⁹ sobre a escolha de alguns "*inputs*" em detrimento de outros no desenho de uma rede neural treinada por inteligência artificial. São eles:

¹⁷ Cetic.br (2018): "Panorama Setorial da Internet", Ano 10, N. 2. Disponível em: https://cetic.br/media/docs/publicacoes/1/Panorama_outubro_2018_online.pdf. Acesso em: 20.02.2019.

¹⁸ Website "This Person Does Not Exist". Disponível em: [ThisPersonDoesNotExist.comhttps://thispersondoesnotexist.com/](https://thispersondoesnotexist.com/). Acesso em 20.02.2019.

¹⁹ "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights". Disponível em: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discriminati_on.pdf. Acesso em 20.02.2019. Cite-se, apenas a título de curiosidade, que a inteligência artificial utilizada para gerar os resultados dito aleatórios (mas que escondem um padrão curioso) é conceituada como uma Generative Adversarial Network (GAN), cujo funcionamento está descrito em: <https://www.lyrn.ai/2018/12/26/a-style-based-generator-architecture-for-generative-adversarial-networks/>. Acesso em: 20.02.2019.

1. Dados mal selecionados, em que os projetistas do sistema algorítmico decidem que certos dados são importantes para a decisão, mas não para outros. No exemplo da “rota mais rápida”, o arquiteto do sistema pode incluir apenas informações sobre estradas, mas não sobre horários de transporte público ou rotas de bicicleta, o que prejudica indivíduos que não possuem um veículo. Tais questões podem ser consideradas como erros qualitativos, em que as escolhas humanas na seleção de certos conjuntos de dados como entradas algorítmicas sobre outras são mal aconselhadas. Escolhas descuidadas de entrada podem levar a resultados tendenciosos - no exemplo da “rota mais rápida”, resultados que podem favorecer rotas para carros, desencorajar o uso do transporte público e criar desertos em trânsito. De forma semelhante, os projetistas podem selecionar dados com muita ou pouca granularidade, resultando em efeitos potencialmente discriminatórios.
2. Dados incompletos, incorretos ou desatualizados, em que pode haver falta de rigor técnico e abrangência na coleta de dados, ou onde inexatidões ou lacunas podem existir nos dados coletados. No exemplo da “rota mais rápida”, isso poderia ocorrer se, por exemplo, o sistema algorítmico não atualizasse as programações de ônibus ou trem regularmente. Mesmo que o sistema funcione perfeitamente em outros aspectos, as instruções resultantes podem desencorajar novamente o uso do transporte público e prejudicar aqueles que não têm alternativas viáveis, como muitos passageiros e residentes de baixa renda.
3. Viés (“bias”) de seleção, por meio do qual o conjunto de entradas de dados para um modelo não é representativo de uma população e, portanto, resulta em conclusões que podem favorecer determinados grupos em detrimento de outros. No exemplo da “rota mais rápida”, se

os dados de velocidade forem coletados apenas dos indivíduos que possuem smartphones, os resultados do sistema podem ser mais precisos para populações mais ricas com concentrações mais altas de telefones inteligentes e menos precisos nas áreas mais pobres onde as concentrações de smartphones são mais baixas.

4. Perpetuação não intencional e promoção de vieses (“bias”) históricos, em que um *loop* de *feedback* faz com que o viés nas entradas ou nos resultados do passado se replique nas saídas de um sistema algorítmico. Por exemplo, quando as empresas enfatizam a “contratação por ajuste cultural” em suas práticas de emprego, elas podem inadvertidamente perpetuar os padrões de contratação passados se a cultura atual do local de trabalho for baseada principalmente em um conjunto específico e restrito de experiências. Em um local de trabalho ocupado principalmente por homens brancos jovens, por exemplo, um sistema algorítmico projetado principalmente para contratar por cultura (sem levar em conta outras metas de contratação, como diversidade de experiência e perspectiva) pode recomendar desproporcionalmente contratar mais homens brancos porque eles pontuam melhor em se encaixar com a cultura.

Assim, como herança de decisões preconceituosas de pessoas ou simplesmente pelos dados refletirem aspectos culturais, históricos e sócio-demográficos existentes na sociedade (Cetic.br, 2018), é importante que haja a expansão de iniciativas de inclusão e diversidade nos setores público e privado para que o perfil dos atores envolvidos no processo de criação e tomada de decisão sobre mecanismos tecnológicos abranja grupos sub representados no ecossistema da tecnologia, pois estes são os principais afetados pela perpetuação do viés e das estruturas que conformam a sociedade contemporânea.

CONCLUSÃO

Este artigo teve como objetivo demonstrar que, a despeito do pouco diálogo que tem sido traçado entre os debates sobre os "perfis" no ecossistema da tecnologia contemporâneo (a identidade virtual dos usuários da rede e o padrão homogêneo dos atores da construção e tomada de decisão sobre mecanismos tecnológicos), ambos são complementares e possuem implicação na perpetuação e manutenção de problemas estruturais e estruturantes da sociedade atual sob a ótica da discriminação.

De um lado, a técnica de "*profiling*" possui riscos associados a práticas de discriminação efetivadas não somente pela coleta de dados pessoais que o sujeito de dados decide compartilhar (primeira camada) mas também pela inferência de padrões de comportamento online e offline (segunda camada) e por tomadas de decisão automatizadas (terceira camada) cujo processo se mantém à margem do escrutínio público e até mesmo da análise do sujeito de dados.

Por outro, percebe-se que, enquanto não houver diversidade e políticas de inclusão no setor privado e público para que a produção de tecnologia seja menos excludente, todos os mecanismos tecnológicos utilizados para uma dita inclusão serão um paradoxo pois reproduzirão os problemas e preconceitos estruturais que compõem e fomentam um processo de retrolegitimação de tais estruturais e, conseqüentemente, de seus preconceitos e discriminações inerentes.

Assim, ainda que as duas questões levantadas pelo presente artigo caminhem paralelamente, ambas terão que ser endereçadas do ponto de vista jurídico, social, científico e ético para que, por meio da intersecção entre elas, se possa determinar que rumo devemos tomar para a construção de um ecossistema tecnológico menos discriminatório, mais inclusivo e mais justo para todos os usuários da rede.

BIBLIOGRAFIA

BKRAN, Maja. AI-Supported Decision-Making under the General Data Protection Regulation.

BROWNSWORD, Roger. Knowing me, knowing you – Profiling, Privacy and the Public Interest, pp. 345 – 363. In: Profiling the European Citizen.

BYGRAVE, L. (2001) Automated Profiling, Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling. *Computer Law & Security Review* 17(1):17-24. COGITAS, Ergo Sum. The Role of Data Protection Law and Non-discrimination Law in Group Profiling in the Private Sector.

BIGONHA, Carolina. (2018). Panorama setorial da Internet: Inteligência Artificial em perspectiva. In: *Inteligência Artificial e ética*, Ano X, Número 2. Cetic.br.

DATATILSYNET, The Norwegian Data Privacy Authority (2018). AI and Privacy.

DEIGHTON, John. "Market Solutions to Privacy Problems?" Chap. 6 in *Digital Anonymity and the Law - Tensions and Dimensions*. Vol. II, edited by C. Nicoll, J.E.J. Prins, and M.J.M. van Dellen, 137 – 146. Information Technology & Law Series. The Hague: T.M.C. Asser Press, 2003.

FEDERAL TRADE COMMISSION OF THE UNITED STATES OF AMERICA (2014). *Data Brokers: A Call for Transparency and Accountability*.

INSTITUTO IGARAPÉ (2018). Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil? Artigo Estratégico 39 | Dezembro 2018.

KAMARINOU, Dimitra et al. "Machine Learning with Personal Data: Profiling, Decisions and the EU General Data Protection Regulation".

KAMBER, Micheline; PEI, Jian. (2012). *Data mining: concepts and techniques*. 3rd ed. ISBN 978-0-12-381479-1.

LETOUZÉ, Emmanuel (2018). Big Data e desenvolvimento: uma visão geral. In: *Big Data para o desenvolvimento*. Ano X, Número 1. Cetic.br.

McCLURG, Andrew J. (2003) A Thousand Words Are Worth a Picture: A Privacy Tort Response to Consumer Data Profiling, 98 Nw. U. L. Rev. 63 (2003).

PASQUALE, Frank. The Black Box Society: The Secret Algorithms That Control Money and Information. ISBN 9780674368279.

PROGRAMARIA; C.A. Technologies. Por que o machismo cria barreiras para as mulheres na tecnologia.

NAVARRO, Ana Maria. N. P. . O Direito Fundamental à Autodeterminação Informativa. In: CONPEDI/UFF. (Org.). XXI Congresso Nacional do CONPEDI. 1ed.: FUNJAB, 2012, v. , p. 429-458.

ROCHA, Elisa Corrêa da. (2015). O direito à autodeterminação informativa e a proteção de dados pessoais.

SABOYA, Maria Clara Lopes (2013): "Relações de Gênero, Ciência e Tecnologia: Uma Revisão da Bibliografia Nacional e Internacional". Educação, Gestão e Sociedade: revista da Faculdade Eça de Queirós, ISSN 2179-9636, Ano 3, número 12, novembro de 2013.

SCARISBRICK-HAUSER, Anne Marie. Data analysis and profiling.

SCHERMER, Bart W (2011). The limits of privacy in automated profiling and data mining. In: *Computer law & security review* 27, pp. 45-52.

SZYMIELEWICZ, Katarzyna. (2019). Sua identidade digital tem três camadas e você só pode proteger uma delas. Tradução e adaptação para o Português por João Carlos Rebello Caribé, com colaboração de Renato Monteiro Leite.

TEIXEIRA, Lucas (2015). Data Brokers e Profiling: vigilância como modelo de negócios.

WHITE HOUSE (2016). "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights", p. 9-10.